

## 数据生产要素视角下开放公共数据与企业创新\*

——基于建立公共数据开放平台的准自然实验

陈艳利<sup>1</sup> 蒋琪<sup>2</sup>

(1. 东北财经大学会计学院/中国内部控制研究中心, 辽宁 大连 116025;

2. 西南政法大学政治与公共管理学院, 重庆 401120)

**内容提要:**畅通要素流通壁垒,推动公共数据作为创新资源向企业流动,鼓励并引导企业利用公共数据实施创新决策和成果转化至关重要。本文以我国各省(自治区/直辖市)陆续开通公共数据开放平台为准自然实验,证实了开放公共数据对企业创新具有促进作用,上述关系以对数据生产要素的资源补给机制和对创新意愿的动力激发机制为双向路径。基于行业性质的细分检验发现,当前释放数据生产要素对实现数字产业化作用明显,但依托产业数字化助力实体经济效果有限。进一步地,本文发现数据传输效率是影响上述关系的关键因素,而现阶段开放文化教育和公共服务性质的数据生产要素效果更佳。与此同时,基于企业既有要素水平的异质性分析发现,对技术、资本等生产要素在位积累性更强的企业而言,上述关系更显著。本文补充了关于数据生产要素实际经济价值的实证证据,为完善数据生产要素流通机制,激励企业释放数据变现价值提供了政策启示。

**关键词:**数据生产要素 公共数据 企业创新 开放公共数据

**中图分类号:**F270 **文献标志码:**A **文章编号:**1002—5766(2024)01—0025—22

## 一、引言

随着通用数字技术的蓬勃发展,数据的大规模可得性使其契合了关键生产要素的核心特质。如同农业经济时代的第一生产要素是土地,工业经济时代的第一生产要素是资本,数据转化为关键生产要素成为经济社会朝着数字经济时代演进的必然结果(于立和王建林,2020)<sup>[1]</sup>。由于数据具有衍生性(Veldkamp,2005)<sup>[2]</sup>,数据的初始获取往往是个体及单位活动踪迹得以提取并使用的结果,这意味着个人、企业和社会团体等都是数据采集的不同主体,公共机构在其中的作用尤为突出(Janssen,2011)<sup>[3]</sup>。公共数据就是国家机关、事业单位、团体组织及科研机构等公共机构在依法履行公共职能过程中生成、采集、记录并保存的数据资源,涉及交通、医疗、就业、文化、卫生、环境、金融、气象等不同领域,是一种规模可观、种类丰富且价值含量极高的数据生产要素。诸多新产品及新公司都在积极利用公共数据,例如,英国的“Mapumental”和德国的“Mapnificent”应用,利用政府提供的地价数据和地方商业登记数据等,为用户提供不动产推荐服务<sup>①</sup>;美国的能源科技公司

收稿日期:2022-11-03

\* 基金项目:国家自然科学基金面上项目“国有资本授权经营:效果评价与机制探索”(72073019);教育部人文社会科学研究规划基金项目“分红监管视角下竞争性国有企业非国有股东治理研究”(20YJA790004);辽宁省教育厅面上项目“数字经济背景下开放公共数据与企业创新——以辽宁省为例”(JYTMS20230636)。

作者简介:陈艳利,女,教授,博士生导师,研究领域为资本市场财务与会计,电子邮箱:chenyanli@dufe.edu.cn;蒋琪,女,博士后,研究领域为公共数据治理,电子邮箱:316299217@qq.com。通讯作者:蒋琪。

① 案例来自开放知识基金会(Open Knowledge Foundation)官网公布的《Open Data Handbook》,详见:<https://opendatahandbook.org>。

OPower利用住宅能源消费调查、人口普查等数据,为客户提供能源使用报告,以节省家庭开支(Thorhildur等,2014)<sup>[4]</sup>;上海知了人力资源有限责任公司利用机动车驾驶证记分、医师执业、护士执业等数据,研发出“知了背调”应用软件,为用人单位提供雇前职业信用背景调查服务。公共数据作为新兴生产要素的经济价值已逐渐在全球范围得到认可。

创新是企业主导下针对生产要素的优化配置与重新组合(夏晗,2022)<sup>[5]</sup>。在当前新旧动能转换以及经济结构转型的关键时期,创新驱动战略成为我国迈向全面现代化的必然路径。新结构经济学认为,创新应以生产要素结构升级为先导,通过新兴生产要素激发闲置要素价值,从而使经济发展与技术进步方向匹配。随着公共数据逐渐满足新兴生产要素条件,利润驱使和竞争压力下,企业可以通过公共数据与传统应用场景的深度融合,寻找新的细分市场并发现提升原有生产力的新方法。但受采集成本高昂、主体集中且行政属性较强等因素影响(Thorhildur等,2014)<sup>[4]</sup>,公共数据在来源上呈现出垄断性,要素流通障碍可能导致资源配置的总体效率及生产要素的利用效率下降(林毅夫和张鹏飞,2006)<sup>[6]</sup>。鉴于此,有序实施公共数据开放才是挖掘数据生产要素创新价值的前提。2014年以来,政府开始积极履行对公共数据的管理与共享责任,围绕企业创新主体地位,推动公共数据加快向企业流动。2021年12月,《关于印发〈要素市场化配置综合改革试点总体方案〉的通知》(国办发[2021]51号)发布,明确提出“建立健全高效的公共数据共享协调机制,支持打造公共数据基础支撑平台,推进公共数据归集整合、有序流通和共享”。2023年2月,国务院印发《数字中国建设整体布局规划》,明确提出“畅通数据资源大循环,推动公共数据汇聚利用”。综合上述背景,探究开放公共数据与企业创新的关系具有重要现实意义。

本文以我国各省(自治区/直辖市)陆续开通公共数据开放平台为准自然实验,实证检验了开放公共数据与企业创新的关系。本文可能存在的创新点在于:第一,补充了关于数据生产要素经济价值的理论分析与经验证据。针对数据生产要素的研究尚有较大进步空间,已有研究缺乏对数据生产要素经济价值的解释力,关于数据生产要素价值演进机制的研究不足,理论认知缺乏系统性分析。本文创新性地构建起开放公共数据与企业创新的理论分析框架,并基于数字经济行业分类、数据质量、数据性质等维度,对上述关系的影响因素、作用机制和异质性展开分析,为证实数据生产要素的实际经济价值提供了理论与实证参考。第二,拓展了开放公共数据经济后果的研究视角。开放公共数据的现有研究多集中在宏观层面,鲜少有研究基于微观企业视角,系统检验开放公共数据进行开发利用和价值创造的效果与路径。同时,研究方法多集中在理论推理、文献回顾、国际比较和案例分析等定性研究中,基于大样本的实证检验相对缺乏,结论客观性和说服力具有一定局限性。本文尝试利用实证数据对开放公共数据进行效果评价,补充了相关领域的实证空缺。第三,丰富了企业创新影响因素的研究结论。本文的研究可纳入企业创新外部影响因素的范畴中。对同领域研究成果而言,将开放公共数据纳入企业创新影响因素范畴,系统性地探讨开放公共数据与企业创新关系的文献还处于缺位状态。数据生产要素、数据创新相较其他生产要素和传统创新形式存在较为显著的差异,传统企业创新的分析框架需要结合数字经济时代背景予以完善,以契合数字时代下经济增长特征的变化要求,本文在一定程度上有所增量贡献。

## 二、文献回顾与理论分析

### 1. 制度变迁视角下开放公共数据的原因

制度是影响社会资源配置效率的重要因素。随着产品和要素相对价格、技术变化 and 市场需求等影响创新的外部因素出现,新的经济机会作为主要激励可能诱发制度变迁。开放公共数据契合了经济社会发展规律和趋势,作为一项新的制度安排,在本质上符合制度变迁的需求动因。首先,数字技术的蓬勃发展为开放公共数据提供支撑前提。卡萝塔·佩蕾丝(2007)<sup>[7]</sup>把大规模可得性视

为关键生产要素的主要特征。第三次信息化浪潮下,随着区块链、人工智能、云计算等数字技术蓬勃发展,数据广泛渗透到全行业、全企业的价值实现环节中,全球数据总量呈现指数型增长,数据的大规模可得性成为现实,具备了成为关键生产要素的基本特征。加之数据处理技术的激进式发展,使得数据收集、储存、管理及分析成本下降,为数据生产要素实现价值转化提供了可能。

其次,降低企业数据生产要素获取成本是开放公共数据的本质需求。从性质上看,公共数据主要包括政府业务数据、社情民情数据和基础环境数据等类型,呈现出体量大、集中度高且辐射范围广等特点。如果企业想依靠自身力量搜集相关数据,难度大且成本高(Thorhildur等,2014)<sup>[4]</sup>:一方面,基础环境数据等数据生产要素需要投入REID标签和读写器、GPS设备、温度湿度传感器等大量物联网感知层设备,及计算、存储和管理数据所需投入的云计算服务等。这类数据生产要素参与生产需要投入高昂的初始成本,一般企业难以承担。另一方面,数据具有衍生性,伴随实际生产活动生成,公共数据是由公共机构在依法履行公共职能等过程中生成、采集、记录并保存的,采集主体集中在国家机关、事业单位、团体组织等公共机构中,这使公共数据在来源上具备垄断性质。鉴于上述原因,政府管理并供给公共数据,成为企业获取此类生产要素的主要甚至唯一途径。与此同时,如果不开放公共数据,让缺乏专业能力的政府直接将这类资源型公共产品转化为制成品,不利于挖掘数据生产要素的内在价值(胡业飞和孙华俊,2021)<sup>[8]</sup>。随着经济生产活动对数据生产要素的利用需求扩张,政府有义务履行对公共数据的管理与共享责任。

最后,数据生产要素内在价值形成的潜在市场是开放公共数据的主要激励。从宏观层面看,当前我国处于经济结构化转型的过渡时期,从数据生产要素到数据创新再到数据资本的转化低于社会稳态水平,基于数据生产要素的投资回报率显著高于边际产出不断下降的传统生产要素,利用数据生产要素赋能传统产业有利于全要素生产率实现倍增。从微观层面看,利用数据生产要素的自身优势进行创新为企业发展带来积极影响,例如数据生产要素的流动性强化了生产过程中的价值创造功能,有利于延长产业链长度(黄贇琳等,2022)<sup>[9]</sup>;数据生产要素的规模报酬递增性有利于实现范围经济(Morton等,2019)<sup>[10]</sup>;数据生产要素的网络效应有助于提高企业的市场敏感性(Bouncken等,2019)<sup>[11]</sup>,引导企业利用长尾效应开辟市场空间并契合顾客需求(Ciampi等,2021)<sup>[12]</sup>。鉴于此,尽可能多地掌握数据生产要素,挖掘数据生产要素内在价值,成为数字经济时代企业发展并建立竞争优势的关键。

## 2. 数据生产要素视角下开放公共数据对企业创新的影响

一方面,畅通公共数据流通壁垒为技术创新注入关键原材料。公共数据在来源上具备政府垄断性质。寻租理论提出,政府部门控制的公共资源会全部构成经济租,成为寻租者重要的租金来源。随着经济生产活动对公共数据的利用需求不断扩张,企业创新迫切需要掌握这种关键性的生产要素。利益驱使下,公共数据如果未能得到有效的公共管理,可能成为数字经济时代企业实施寻租活动的重要目标。企业创新是一项资源消耗性活动,具有很强的资源依赖性。如果不对公共数据实施有效的开放共享,要素流通障碍可能导致这项关键创新资源配置不力,这反而会加大企业实施公共数据创新的基础成本,造成社会整体资源浪费。寻租理论提出,加强要素流通有利于消除生产要素在产业中形成的“租”。随着国家对数字经济发展战略部署的不断完善,政府建立起公共数据开放平台,形成了连接公共数据发布端与获取端的中介和载体。这从根本上促进了公共数据的有效流通,实现公共数据作为创新要素向企业聚集。

随着这种生产要素进入市场和企业,产业结构和企业经营随之发生深刻变化。从宏观视角看,公共数据作为高效率生产要素进入市场,能够促进原有的生产要素配置结构升级,使之匹配技术进步方向。新结构经济学提出,经济发展需要以要素结构升级为路径。技术选择假说认为,一旦要素禀赋实现升级,利润与竞争压力将会推动企业自发进行技术创新与产业升级(林毅夫和张

鹏飞, 2006)<sup>[6]</sup>。因而技术进步偏向效率更高的生产要素, 要素配置应向高效率生产要素倾斜(Basu 和 Weil, 1998)<sup>[13]</sup>。数据不同于资本、劳动等传统生产要素, 其自由流动性和规模报酬递增性等特征有利于提高要素投入的增长幅度, 是一种高效率的新生产要素。随着这种高效率的生产要素进入生产, 要素禀赋结构和技术进步方向会在利润与竞争压力的驱使下, 自然地与要素效率水平相匹配, 从而促进经济实现创新发展。

从微观视角看, 公共数据作为新兴生产要素更新原有生产函数并形成创新“新组合”。熊彼特创新理论以“创新”和“新组合”来解释经济发展, 认为创新在本质上是企业对生产要素的重新组合(约瑟夫·熊彼特, 2009)<sup>[14]</sup>。随着公共数据作为新生产要素进入现有生产函数, 为创新开发者提供了前瞻性的研究边界与科学性的研究工具(Einav 和 Levin, 2014)<sup>[15]</sup>, 以数据生产力为驱动力, 促进研发模式转型的同时, 带动企业内部组织构成、经营管理和资源配置等全面转变, 更新商业模式并拓展企业为利益相关者提供的价值实现路径(Ciampi 等, 2021)<sup>[12]</sup>。从本质上讲, 这些改变就是既有生产函数中的要素构成与结构关系在发生变化(史丹, 2022)<sup>[16]</sup>, 产业模式、商业形态和管理流程衍生出“新组合”的过程, 就是企业创新的过程。

另一方面, 释放公共数据有利于激发企业创新的内在动力。首先, 按照熊彼特创新理论的观点, 利润是技术创新与扩散的驱动力。数据生产要素具有正外部性, 利用数据生产要素实现的创新成果普遍易于高效传播与复制(Ciriello 等, 2018)<sup>[17]</sup>, 这种自生长性进一步强化了数字创新成果的持续迭代需求和实现能力, 进而提高了社会创新频率与竞争程度。随着公共数据的适配场景和潜在功能逐渐被发现并传播, 市场领头者已实现的创新方案能够发挥规模性的示范与宣传作用, 降低创新活动的不确定性, 并通过潜在利润吸引更多后继者。

其次, 数据生产要素具有非损耗性, 折旧率低, 边际成本趋近于零。数据生产要素突破了有限资源对经济发展的局限, 边际收益呈现递增。唐家等(2022)<sup>[18]</sup>发现, 从长期来看, 数字经济发展和高集中度市场会通过彼此强化形成市场垄断趋势, 数字经济时代的经济发展具备典型的“赢者通吃”定律(Guellec 和 Paunov, 2017)<sup>[19]</sup>, 通过持续数据创新来应对挑战并保持优势成为企业的不二选择(Hagiu 和 Wright, 2015)<sup>[20]</sup>。面对激烈的市场竞争, 有利于激发企业家精神, 鼓励企业积极迎接数字经济时代的挑战, 充分挖掘新生产要素价值并实施创新活动。

最后, 交易费用对企业的生产决策与资源配置有着重要影响, 威廉姆森把交易不确定性、交易频率和竞争对手数量等作为影响交易费用的关键因素。随着利用公共数据进行创新交易频率提高, 越来越多的企业加入竞争, 扩大了数据应用场景并完善了数据应用方案, 这极大地降低了潜在进入者实施创新的风险不确定性。与此同时, 对数据生产要素的创新应用能够改善企业自身信息传递效率(肖土盛等, 2022)<sup>[21]</sup>, 这有利于提高创新资源配置与利用效率。王超贤等(2022)<sup>[22]</sup>提出, 利用数据生产要素进行价值创造, 在本质上就是通过数据来优化决策。随着数据生产要素对企业组织结构不断改造, 企业在经营管理过程中的决策有效性提高, 因信息不对称、有限理性和机会主义等局限带来的信息搜寻与决策成本得以降低。数据生产要素发挥决策支撑作用, 有利于实施柔性化和精细化的流程管理, 提升创新活动与市场需求的匹配效率, 减少供需缺口, 降低创新风险。

因此, 本文提出如下假设:

H<sub>1</sub>: 开放公共数据能够促进企业创新。

### 3. 数据质量对开放公共数据与企业创新关系的影响

肖土盛等(2022)<sup>[21]</sup>认为, 数据生产要素的价值是极大量、多种类、速变动数据的关联与合成, 当具备上述特征时, 数据生产要素更容易实现价值创造。因此, 数据生产要素的价值受质量属性影响显著, 高质量的数据生产要素被普遍认为与其他生产要素的融合效率更高(林志杰和孟政炫,

2021)<sup>[23]</sup>。本文认为,数据质量对开放公共数据与企业创新的关系具有正向调节作用:首先,开放公共数据的总量越大,越有利于形成绝对生产力,产生规模经济效应,降低创新的单位成本。从根源上看,数据生产要素发挥作用的基础在于充足的数据资源聚集,只有当数据生产要素具备一定规模时,才能产生持续积累的绝对生产力。从成本上看,数据生产要素初始成本高,但边际成本却低。这意味着后续投入越多的数据生产要素参与到再生产环节,前期投入的专用性资产越能得以充分摊销,单位使用成本才能随之降低,扩大企业边界并激发创新意愿。与此同时,数据生产要素与数字技术不可分割,数据生产要素作为原材料投入越多,在数据衍生性的影响下,会形成滚雪球优势,放大数字技术的外溢性,这有利于发挥数据生产要素对企业生产经营全面渗透的驱动力,进一步刺激规模经济效应。

其次,开放公共数据的多样性越丰富,越有利于形成网络联动,激发范围经济效应,形成刺激创新的需求激励。利用数据生产要素实现的创新产品一般通过应用场景对生产经营和社会活动发挥渗透作用。数据生产要素具有正外部性,在梅特卡夫效应的作用下,其种类越丰富,可供形成关联的应用场景就越多,扩散效应也越大(Nambisan等,2017)<sup>[24]</sup>。这些场景化的数据应用有助于形成一张强关系的多维度数据网络,衍生一系列“数据+”类型的生产能力,逐渐将社会中的每一个个体和组织卷入数字经济时代,通过数据网络效应形成数字孪生,并由此分化为一系列新的产品、服务和组织形式。实体空间在数据孪生的持续扩张下,场景取代功能,强关系取代弱关系,动态网络取代静态网络,顾客对以数据生产要素作为决策依据与行动指引的依赖程度不断提高,市场需求得以巩固。

最后,开放的公共数据传输效率越高,越有利于提升要素协同效率,发挥长尾效应,降低企业创新风险。数据生产要素与其他生产要素相比,具有高流动性。生产要素的流动性越高,越容易形成聚集、积累和乘数效应,协同提升流动性弱的其他生产要素配置效率(刘玉奇和王强,2019)<sup>[25]</sup>。因而,当公共数据开放平台提供的公共数据传输效率越高,更新越及时,公共数据与企业原始数据、技术和设备的复用性、扩展性和耦合性越高。与此同时,数据生产要素的传输效率越高,在挖掘需求、识别机会和化解风险等方面的反馈越灵敏,这不仅有利于助力企业捕捉市场需求并判断风险,还有利于降低信息搜寻成本,结合个性化需求利用长尾效应提升产品质量,从而提升数据创新成果的持续迭代需求和实现能力。

基于上述分析,本文提出如下假设:

H<sub>2</sub>:数据质量越高,开放公共数据对企业创新的促进作用越显著。具体来看,当开放的公共数据总量越大、多样性越丰富、传输效率越高时,开放公共数据对企业创新的促进作用越显著。

#### 4. 数据性质对开放公共数据与企业创新关系的影响

数据生产要素能否创造经济价值取决于两个方面:第一,市场主体是否具有利用数据生产要素的价值倾向;第二,传统的经济生产方式是否具备与数据生产要素融合的基础(王开科等,2020)<sup>[26]</sup>。结合上述两点,在考虑数据性质对开放公共数据与企业创新关系的影响时,本文做出如下分析:一方面,开放与市场需求关联度更高的公共数据,企业利用公共数据进行创新的动力越足。中国是从消费端开启的数字化。刘淑春(2019)<sup>[27]</sup>提出,我国的数字经济呈现出“三二一”产业逆向渗透趋势。随着新零售、新服务等经济业态不断涌现,购物、餐饮、娱乐、教育、医疗和出行等各类数字经济应用场景得以开辟,并不断向产业上游传导,形成数字化流量入口。姜松和孙玉鑫(2020)<sup>[28]</sup>发现,数字经济对实体经济产生了一定的挤出效应,目前教育、金融等服务型产业与数据生产要素和数字经济的融合程度相对更高。因此,服务业等具有膨胀市场需求的行业,其利润与竞争优势更能吸引企业竞相参与创新。

另一方面,如果与公共数据性质呈现高关联度的行业数字经济发展基础越好,企业利用公共

数据进行创新的难度越小。对数字经济发展基础较好的行业而言,市场可供评估及参考的创新应用方案越丰富,这有利于降低企业利用公共数据实施创新的信息搜寻成本,规避结果的不确定性,从而提振企业的创新动力。数据生产要素的长尾效应使创新活动重心由生产转向服务,因而数字经济在服务业的渗透率相对最高。事实上,上述观点也在实践中得到了检验。当前聚焦在提升居民消费质量和办事效率等服务需求的数据应用方案较为成熟,各种餐饮、旅游、导航、购物和政务服务等平台软件及专业社区层出不穷,容易形成模仿效应带动后继企业加入。换言之,如果开放的公共数据性质与这些领域更为接近,企业实施创新的难度相对更低。

因此,本文提出如下假设:

H<sub>3</sub>:数据性质会对开放公共数据对企业创新的促进作用产生影响,当开放的公共数据与市场需求关联度越高或涉及行业的数字经济发展基础越好时,开放公共数据对企业创新的促进作用越显著。

### 三、研究设计

#### 1. 模型设计

自2012年北京和上海两地率先实施公共数据开放以来,截至2021年我国已有25个省级(含直辖市/自治区)政府建立了公共数据开放平台(如表1所示)。公共数据开放平台在政府推进公共数据开放的过程中发挥着核心作用,整合了各公共部门的分散数据,成为公共数据发布端和获取端的中介与载体。本文拟选取A股上市公司作为研究样本,以我国各省(自治区/直辖市)陆续开通并运营公共数据开放平台为准自然实验,构建多期双重差分模型<sup>①</sup>。

表1 2012—2021年各省(自治区/直辖市)公共数据开放平台上线概况

年份	省份/自治区/直辖市
2012年	北京市、上海市
2015年	浙江省
2016年	广东省、贵州省
2018年	河南省、江西省、宁夏回族自治区等
2019年	新疆维吾尔自治区、海南省、四川省等
2020—2021年	广西壮族自治区、湖北省、湖南省等

$$Innovation_{i,t} = \beta_0 + \beta_1 Reform_{i,t} + \sum controls_{i,t} + Firm\_F.E + Year\_F.E + Province\_F.E + \epsilon_{i,t} \quad (1)$$

$$Innovation_{i,t} = \beta_0 + \beta_1 Reform_{i,t} + \beta_2 Character_{i,t} + \beta_3 Character_{i,t} \times Reform_{i,t} + \sum controls_{i,t} + Firm\_F.E + Year\_F.E + Province\_F.E + \epsilon_{i,t} \quad (2)$$

$$Innovation_{i,t} = \beta_0 + \beta_1 Reform_{i,t} + \beta_2 Category_{i,t} + \beta_3 Category_{i,t} \times Reform_{i,t} + \sum controls_{i,t} + Firm\_F.E + Year\_F.E + Province\_F.E + \epsilon_{i,t} \quad (3)$$

其中,式(1)~式(3)中的  $Innovation_{i,t}$  表示  $i$  公司在第  $t$  年的创新水平。 $Reform_{i,t}$  为哑变量,表示  $i$  公司在第  $t$  年时,其所在省(自治区/直辖市)是否开通了公共数据开放平台,该变量是模型的政策效应观测变量。 $\sum controls_{i,t}$  表示  $i$  公司第  $t$  年的控制变量合集。 $Firm\_F.E$ 、 $Year\_F.E$  和  $Province\_F.E$  分别表示公司、年份和省份固定效应。 $\epsilon_{i,t}$  代表随机扰动项。同时,该模型还在公司层面对标准误进

<sup>①</sup> 需要说明的是:第一,参考《中国地方政府数据开放报告》的评价标准,本文所指的公共数据开放平台需具备 gov.cn 的官方平台域名,形式为统一专有式(即数据统一公开在一个专门平台),且平台确实开放了电子格式、可供下载或通过 API 接口获取的、结构化的数据集等特征。第二,部分省份尽管在报告或新闻中曾公布过已建立公共数据开放平台,但在研究期间进入官网后出现了 404 错误、505 错误或显示“无法访问此界面”等类似情况,本文将这些省份视同归属对照组。

行了聚类修正。式(1)用于检验假设  $H_1, \beta_1$  为结果观察项,如果开放公共数据对企业创新具有促进作用,其预计符号为正。式(2)用于检验假设  $H_2, Character_{i,t}$  表示  $i$  公司所在省(自治区/直辖市)第  $t$  年开放公共数据的数据质量。式(3)用于检验假设  $H_3, Category_{i,t}$  表示  $i$  公司所在省(自治区/直辖市)第  $t$  年开放公共数据的数据性质。观测结果时,如果式(2)和式(3)中的  $\beta_3$  显著且与  $\beta_1$  方向一致,那么该项数据质量或性质会对开放公共数据与企业创新的关系具有促进作用。

## 2. 变量设定

(1)被解释变量。已有研究通常按照创新投入和创新产出两类视角评价企业创新水平(徐悦等,2018)<sup>[29]</sup>。本文从创新投入视角出发,参考徐悦等(2018)<sup>[29]</sup>、郭玥(2018)<sup>[30]</sup>等做法,将企业研发支出除以营业收入的结果命名为 *Innovation* 用于检验。

(2)解释变量。*Reform* 是模型的政策效应观测项,其为哑变量,表示  $i$  公司在第  $t$  年,其所在省(自治区/直辖市)是否开通了公共数据开放平台,如果是则取 1,否则取 0。

(3)控制变量。参考解维敏等(2009)<sup>[31]</sup>、李春涛等(2020)<sup>[32]</sup>等研究,本文分别从经营状况、财务状况、公司治理和地区发展等方面控制了其他控制变量,具体情况如表 2 所示。

(4)调节变量。数据质量(*Character*)由三个维度的指标构成:一是数据总量(*Amount*),利用  $i$  公司第  $t$  年所在省(自治区/直辖市)公共数据开放平台共享的公共数据总条数进行度量;二是数据多样性(*Sources*),数据多样性利用  $i$  公司第  $t$  年所在省(自治区/直辖市)开放的公共数据来源的部门总数进行度量;三是数据传输效率(*Quality*),利用  $i$  公司第  $t$  年所在省(自治区/直辖市)开放的通过 API 接口<sup>①</sup>传输的公共数据总条数进行度量。上述变量除以地区经济生产总值进行标准化处理。对数据性质(*Category*)而言,各省(自治区/直辖市)公共数据开放平台已实现分类标签共计 49 类。通过归类与合并,本文将这些平台发布的数据生产要素性质分为六类,并相应统计了各平台不同性质数据生产要素的条数。本文利用  $i$  公司所在省(自治区/直辖市)第  $t$  年不同性质数据生产要素的数量,除以地区经济生产总值进行度量。相关变量的分类和度量如表 2 和表 3 所示。

表 2 变量定义

变量类型	变量符号	变量名称	计算方法
被解释变量	<i>Innovation</i>	企业创新	研发支出/营业收入
解释变量	<i>Reform</i>	政策效应观测项	表示 $i$ 企业在 $t$ 时期时,其所在省(自治区/直辖市)是否开通了公共数据开放平台,如果开通了则取 1,否则取 0
控制变量	<i>ROA</i>	总资产收益率	净利润/总资产
	<i>Size</i>	企业规模	ln(年末总资产)
	<i>Lev</i>	资产负债率	(短期借款+长期借款+应付债券)/资产总计
	<i>Capital</i>	资本密集率	资产总计/营业收入
	<i>TAI</i>	有形资产密集度	有形资产总额/资产总计,有形净资产总额=资产总计-无形资产净值-商誉净额
	<i>Growth</i>	销售增长率	(本期销售收入-上期销售收入)/上期销售收入
	<i>IAR</i>	无形资产比率	无形资产/资产总计
	<i>Board</i>	董事会规模	ln(董事会人数)
	<i>Manager</i>	高管持股比例	高级管理人员持股数量/总股数
	<i>IDR</i>	独立董事占比	独立董事人数/董事人数

① API 即应用程序接口(application program interface),是一组定义、程序及协议的集合,可实现计算机软件之间的直接相互通信。通过 API 接口可以使数据获取更加容易,开发人员可以通过 API 接口程序直接开发应用程序,减少编写无用程序,减轻编程任务,因此通过 API 接口传输的数据相对质量更高。

续表 2

变量类型	变量符号	变量名称		计算方法	
控制变量	<i>GDP</i>	地区生产总值		地区生产总值(亿元)/地区个体及企业数量总和(户)	
	<i>Finance</i>	地区财政收入		一般公共预算收入(亿元)/地区个体及企业数量总和(户)	
	<i>Tax</i>	地区税收收入		税收收入(亿元)/地区个体及企业数量总和(户)	
调节变量	<i>Character</i>	数据质量	<i>Amount</i>	数据总量	<i>i</i> 公司第 <i>t</i> 年所在省(自治区/直辖市)公共数据开放平台共享的公共数据总条数/地区经济生产总值(亿元)
			<i>Sources</i>	数据多样性	<i>i</i> 公司第 <i>t</i> 年所在省(自治区/直辖市)开放的公共数据来源的部门总数/地区经济生产总值(亿元)
			<i>Quality</i>	数据传输效率	<i>i</i> 公司第 <i>t</i> 年所在省(自治区/直辖市)开放的通过 API 接口传输的公共数据总条数/地区经济生产总值(亿元)
	<i>Category</i>	数据性质		<i>i</i> 公司第 <i>t</i> 年所在省(自治区/直辖市)公共数据开放平台共享的某类性质公共数据总条数/地区经济生产总值(亿元)	

表 3 数据性质分类

指标分类	指标名称
经济生产	<i>Economic</i>
社会民生	<i>Society</i>
文化教育	<i>Cultural</i>
资源能源	<i>Resources</i>
公共服务	<i>Public</i>
其他类	<i>Others</i>

### 3. 样本选择与数据来源

本文选取 A 股上市公司作为研究样本,并根据对上市公司所在省(自治区/直辖市)是否开通了公共数据开放平台,进一步划分为实验组和对照组。样本期间设置为 2007—2020 年,主要的微观层面数据来源于 CSMAR 数据库,宏观层面数据来源于《中国统计年鉴》。各省(自治区/直辖市)开通公共数据开放平台的情况,通过政府报告、门户网站、新闻报道等渠道手工整理得到。为确保研究结论有效,本文剔除了数据缺失类、金融类以及 ST、\*ST 类样本,并对所有连续变量在 1% 分位数处进行了缩尾处理。本文的数据处理软件为 Stata 15.0 和 Excel 2016。

## 四、实证分析

### 1. 描述性统计分析

本文最终筛选出 20219 家公司-年样本,表 4 为描述性统计分析结果。从被解释变量看,企业创新(*Innovation*)的平均值为 0.044,最大值为 0.257,最小值为 0,说明企业创新水平存在较为显著的差异。政策效应观测项(*Reform*)的均值是 0.426,表示 2007—2020 年,42.6% 的上市公司所在地陆续建立了省(自治区/直辖市)级公共数据开放平台。其他控制变量的描述性统计分析结果与既有文献基本一致,不再赘述。鉴于数据质量(*Character*)相关变量均按地区经济生产总值进行过标准化处理,以 2022 年我国各省 GDP 收入约 39000 亿元<sup>①</sup>的均值推算,开通了公共数据开放平台的省份共计释放公共数据总条数的均值为 2145 条,涉及公共部门数量的均值为 39 个,通过 API 接口传输的公共数据总条数均值为 1755 条。从最大值和最小值差距上看,数据总量(*Amount*)和数据传输效率(*Quality*)方面,各省(自治区/直辖市)差距相对较大,数据多样性(*Sources*)差距相对较小。至

<sup>①</sup> 数据来自国家统计局发布的报告《2022 年国民经济顶住压力再上新台阶》,详见: [https://www.gov.cn/xinwen/2023-01/17/content\\_5737453.htm](https://www.gov.cn/xinwen/2023-01/17/content_5737453.htm)。报告显示,2022 年全年国内生产总值为 1210207 亿元,故各省生产总值均值为:1210207÷31=39039 亿元。

于数据性质 (*Category*) 维度, 目前对经济生产 (*Economic*) 和公共服务 (*Public*) 领域的公共数据共享最多。与此同时, 归属其他类 (*Others*) 的公共数据同样占有一定份额, 表明在数据分类工作上还需进一步做实做细。

表 4 描述性统计分析

变量	样本数	平均值	标准差	最小值	中位数	最大值
<i>Innovation</i>	20219	0.044	0.044	0.000	0.035	0.257
<i>Reform</i>	20219	0.426	0.495	0.000	0.000	1.000
<i>ROA</i>	20219	0.048	0.059	-0.276	0.046	0.196
<i>Size</i>	20219	22.090	1.283	19.880	21.890	26.050
<i>Lev</i>	20219	0.392	0.200	0.048	0.379	0.898
<i>Capital</i>	20219	2.192	1.453	0.423	1.818	9.534
<i>TAI</i>	20219	0.923	0.088	0.541	0.953	1.000
<i>Growth</i>	20219	0.279	0.651	-0.643	0.121	4.342
<i>IAR</i>	20219	0.047	0.044	0.000	0.036	0.264
<i>Board</i>	20219	2.130	0.196	1.609	2.197	2.708
<i>Manager</i>	20219	0.085	0.150	0.000	0.003	0.622
<i>IDR</i>	20219	0.375	0.053	0.333	0.333	0.571
<i>GDP</i>	20219	0.008	0.004	0.002	0.009	0.019
<i>Finance</i>	20219	0.054	0.311	0.000	0.001	2.149
<i>Tax</i>	20219	0.044	0.253	0.000	0.001	1.782
<i>Amount</i>	20219	0.055	0.097	0.000	0.014	0.500
<i>Sources</i>	20219	0.001	0.002	0.000	0.001	0.009
<i>Quality</i>	20219	0.045	0.091	0.000	0.001	0.452
<i>Economic</i>	20219	0.015	0.028	0.000	0.004	0.187
<i>Society</i>	20219	0.008	0.013	0.000	0.003	0.066
<i>Cultural</i>	20219	0.005	0.007	0.000	0.001	0.034
<i>Resources</i>	20219	0.004	0.005	0.000	0.002	0.030
<i>Public</i>	20219	0.021	0.057	0.000	0.001	0.299
<i>Others</i>	20219	0.001	0.004	0.000	0.000	0.030

## 2. 主要回归结果

(1) 基准回归结果。表 5 列示了开放公共数据对企业创新影响的实证结果。第 (2) 列是在第 (1) 列控制公司和年份层面固定效应的基础上, 增加了省份层面的固定效应, *Reform* 均在 5% 的基础上正向显著, 印证了研究假设  $H_1$ , 即开放公共数据能够促进企业创新。

表 5 基准回归结果

变量	(1)	(2)
<i>Reform</i>	0.001** (2.06)	0.001** (2.02)
<i>ROA</i>	-0.040*** (-6.41)	-0.039*** (-6.37)
<i>Size</i>	-0.001 (-1.26)	-0.001 (-1.14)
<i>Lev</i>	-0.023*** (-7.82)	-0.023*** (-7.85)

续表 5

变量	(1)	(2)
<i>Capital</i>	0.006*** (9.80)	0.006*** (10.15)
<i>TAI</i>	0.019*** (2.68)	0.019*** (2.73)
<i>Growth</i>	0.000 (0.06)	0.000 (0.09)
<i>IAR</i>	0.048*** (3.32)	0.052*** (3.57)
<i>Board</i>	0.006** (2.23)	0.006** (2.19)
<i>manager</i>	0.004 (1.15)	0.005 (1.20)
<i>IDR</i>	-0.004 (-0.54)	-0.003 (-0.47)
<i>GDP</i>	0.144 (0.73)	0.126 (0.63)
<i>Finance</i>	0.008 (0.92)	0.007 (0.77)
<i>Tax</i>	-0.007 (-0.81)	-0.005 (-0.62)
年份固定效应	是	是
公司固定效应	是	是
省份固定效应	否	是
常数项	0.009 (0.41)	0.004 (0.17)
观测值	20219	20219
R <sup>2</sup>	0.177	0.182

注: \*、\*\*、\*\*\*表示 10%、5%、1% 的显著性水平,下同

(2)基于数据质量的回归结果。以数据质量作为调节效应的回归结果如表 6 所示,数据传输效率与政策效应观测项的交乘项显著,且与政策效应观测项的方向一致,验证了研究假设 H<sub>2</sub>。但数据总量和数据多样性的检验结果却在预期之外。结合利用公共数据实施创新的客观现状,本文尝试分析其中原因:首先,数据总量并未显示出正向调节作用,这可能是因为单纯依靠量上的提升无法对创新产生实质性影响,真正有价值的原材料才符合生产要素本质。换言之,开放公共数据在总量上的提升应建立在保证质量的前提下。其次,数据多样性并未显示正向调节作用,可能是因为当前对于公共数据的创新利用尚处于实践初期,数据利用往往局限于单一类型,企业对关联性应用场景的拓展不足,范围经济效应有待深化。

表 6 数据质量、开放公共数据与企业创新

变量	(1)	(2)	(3)
	数据总量	数据多样性	数据传输效率
<i>Reform</i>	0.001** (1.99)	0.002** (2.14)	0.001** (1.96)

续表 6

变量	(1)	(2)	(3)
	数据总量	数据多样性	数据传输效率
<i>Amount</i>	0.009 (0.58)		
<i>Reform</i> × <i>Amount</i>	0.016 (1.60)		
<i>Sources</i>		0.697 (0.86)	
<i>Reform</i> × <i>Sources</i>		0.641 (1.21)	
<i>Quality</i>			0.015 (0.83)
<i>Reform</i> × <i>Quality</i>			0.017* (1.83)
控制变量	控制	控制	控制
年份固定效应	是	是	是
公司固定效应	是	是	是
省份固定效应	是	是	是
常数项	0.000 (0.01)	0.001 (0.05)	0.001 (0.02)
观测值	20219	20219	20219
R <sup>2</sup>	0.183	0.182	0.183

(3)基于数据性质的回归结果。本文以公共数据的一级指标分类情况作为数据性质的分类标准,回归结果如表 7 所示。由表 7 可知,文化教育和公共服务与政策效应观测项的交乘项显著,且与政策效应观测项的方向一致,说明该类数据性质对开放公共数据与企业创新的关系具有正向调节作用。综合来看,这两类性质的公共数据与市场需求关联度较大,与数字经济应用场景较为匹配,与假设 H<sub>3</sub>基本契合。

表 7 数据性质、开放公共数据与企业创新

变量	(1)	(2)	(3)	(4)	(5)	(6)
	经济生产类	社会民生类	文化教育类	资源能源类	公共服务类	其他类
<i>Reform</i>	0.001 (1.64)	0.001* (1.87)	0.001** (2.08)	0.001* (1.93)	0.001* (1.92)	0.002** (2.06)
<i>Economic</i>	-0.102* (-1.73)					
<i>Reform</i> × <i>Economic</i>	-0.023 (-0.71)					
<i>Society</i>		-0.001 (-0.01)				
<i>Reform</i> × <i>Society</i>		0.094 (1.38)				
<i>Cultural</i>			0.322 (1.49)			
<i>Reform</i> × <i>Cultural</i>			0.266** (2.17)			

续表 7

变量	(1)	(2)	(3)	(4)	(5)	(6)
	经济生产类	社会民生类	文化教育类	资源能源类	公共服务类	其他类
<i>Resources</i>				0.067 (0.25)		
<i>Reform</i> × <i>Resources</i>				0.146 (0.86)		
<i>Public</i>					0.010 (0.38)	
<i>Reform</i> × <i>Public</i>					0.027* (1.78)	
<i>Others</i>						-0.013 (-0.03)
<i>Reform</i> × <i>Others</i>						-0.153 (-0.76)
控制变量	控制	控制	控制	控制	控制	控制
年份固定效应	是	是	是	是	是	是
公司固定效应	是	是	是	是	是	是
省份固定效应	是	是	是	是	是	是
常数项	0.007 (0.31)	0.000 (0.01)	-0.004 (-0.17)	0.002 (0.08)	0.002 (0.09)	0.004 (0.16)
观测值	20219	20219	20219	20219	20219	20219
R <sup>2</sup>	0.183	0.183	0.183	0.182	0.183	0.182

### 3. 稳健性检验

(1) 替换被解释变量。稳健性检验中,本文从创新投入视角补充了 *Innovation\_A*,即经行业年度均值调整后的 *Innovation* 进行测试。参考鞠晓生等(2013)<sup>[33]</sup>做法,基于创新产出视角的指标(*Innovation\_B*),即无形资产增量除以资产总计作为企业创新产出的代理变量。同时,参考陶锋等(2023)<sup>[34]</sup>做法,构建出数字技术创新专利衡量指标(*Innovation\_C*),具体计算时取数字经济专利申请总数加1的自然对数。该指标相比直接的专利申请数据,能从技术领域对数据创新产出情况提供更加贴切的度量。鉴于创新产出存在一定滞后性,本文还利用 *i* 公司第 *t*+1 期的数字技术创新专利加1取自然对数构造(*Innovation\_D*)用于稳健性检验。使用 *Innovation\_A*、*Innovation\_B*、*Innovation\_C* 和 *Innovation\_D* 替换原被解释变量,由表 8 可知,*Reform* 均在 5% 的水平上显著,且方向与预期相符,结果稳健。

表 8 稳健性检验汇总表 1

变量	(1)	(2)	(3)	(4)
	<i>Innovation_A</i>	<i>Innovation_B</i>	<i>Innovation_C</i>	<i>Innovation_D</i>
<i>Reform</i>	0.001** (2.03)	0.001** (2.30)	0.063** (2.28)	0.065** (2.35)
控制变量	控制	控制	控制	控制
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是
省份固定效应	是	是	是	是
常数项	-0.033 (-1.46)	-0.066*** (-4.66)	-2.727*** (-2.85)	-1.688* (-1.74)
观测值	20219	18668	15615	15402
R <sup>2</sup>	0.181	0.225	0.092	0.082

(2)平行趋势检验。满足平行趋势假设是应用双重差分法的前提。本文借鉴 Beck 和 Levkov (2010)<sup>[35]</sup>做法,将实验组样本期拆分为不同期间,以便观察开放公共数据对企业创新在不同阶段的影响。式(4)中, $pre\_m$ 、 $aft\_n$ 、 $current$ 均为虚拟变量,如果样本属于实验组且处于政策实施前第  $m$  期,则  $pre\_m$  取值为 1,否则为 0;如果样本属于实验组且处于政策实施后第  $n$  期,则  $aft\_n$  取值为 1,否则为 0;如果样本属于实验组且处于政策实施当期,则  $current$  取值为 1,否则为 0。回归结果如图 1 所示, $pre\_5\sim pre\_1$  的置信区间均穿过 0 轴,满足了平行趋势假设。

$$\begin{aligned}
 Innovation_{i,t} = & \beta_0 + \beta_1 pre\_5 + \beta_2 pre\_4 + \beta_3 pre\_3 + \beta_4 pre\_2 + \beta_5 pre\_1 + \beta_6 current \\
 & + \beta_7 aft\_1 + \beta_8 aft\_2 + \beta_9 aft\_3 + \beta_{10} aft\_4 + \beta_{11} aft\_5 + \sum controls_{i,t} \\
 & + Firm\_F.E + Year\_F.E + Province\_F.E + \epsilon_{i,t}
 \end{aligned} \tag{4}$$

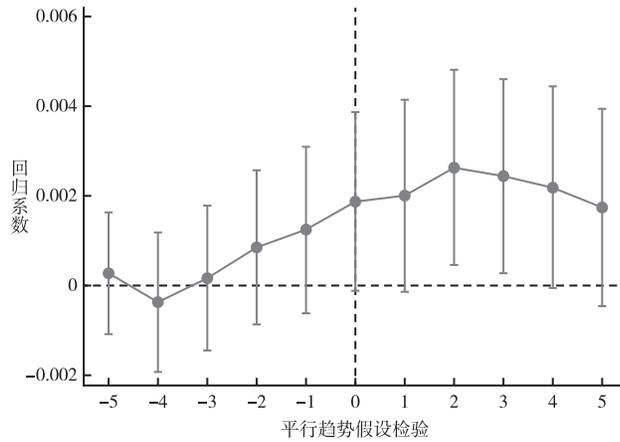


图 1 平行趋势检验

(3)倾向性得分匹配法。本文以主回归模型中的连续型控制变量为协变量,对实验组进行了 Logit 回归,并根据预测得分通过最近邻匹配法以 0.05 为卡尺门槛,按照 1:2 规则匹配实验组和对照组,匹配后共计 14621 个样本符合要求,回归结果如表 9 第(1)列所示,其回归结果仍旧显著,结果稳健。

表 9 稳健性检验汇总表 2

变量	(1)	(2)	(3)	(4)	(5)
	PSM	剔除 2018 年及以后样本	剔除上海市样本	剔除贵州省样本	调整样本结构
<i>Reform</i>	0.002** (2.32)	0.002** (2.13)	0.001* (1.78)	0.002** (2.07)	0.001** (2.16)
控制变量	控制	控制	控制	控制	控制
年份固定效应	是	是	是	是	是
公司固定效应	是	是	是	是	是
省份固定效应	是	是	是	是	是
常数项	0.017 (0.63)	0.047* (1.83)	0.032 (1.30)	0.003 (0.12)	0.022 (1.30)
观测值	14621	12167	18781	20034	11877
R <sup>2</sup>	0.174	0.166	0.183	0.182	0.164

(4)安慰剂测试。鉴于被解释变量的统计显著性可能源于某些难以观察的随机因素,本文随机打乱实验组与对照组分组情况进行安慰剂测试。首先,利用计算机随机生成新的处理组,将上述过程重复进行 1000 次回归;然后,画出所有回归结果中  $t$  值的核密度图;最后,选择表 5 第(2)列

*Reform* 的  $t$  值进行对比。由图 2 可知,仅有极少数回归的  $t$  值大于表 5 第(2)列实际回归的  $t$  值,结果仍旧稳健。

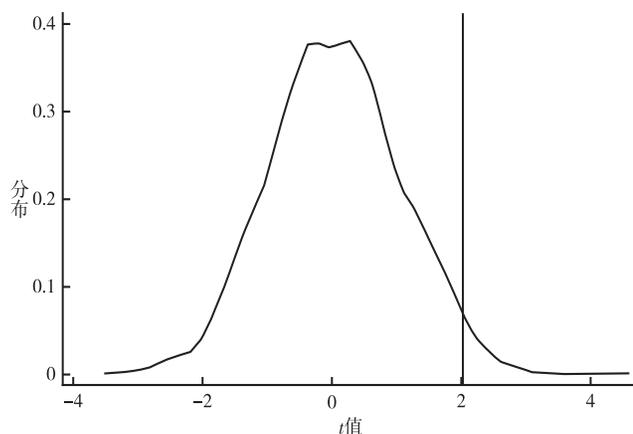


图 2 安慰剂测试——随机抽样回归结果的  $t$  值核密度图

(5)排除同期其他政策干扰。2018年1月发布的《公共信息资源开放试点工作方案》将北京、上海、浙江、福建和贵州五地作为公共信息资源开放试点。鉴于该试点改革与本文的研究期间存在同期性,为排除干扰,本文剔除了2018年及其以后的样本,回归结果如表9第(2)列所示。由表9可知,在剥离公共信息资源开放试点工作的影响后,回归结果仍旧显著,结果稳健。2012年,我国颁布了《国家智慧城市试点暂行管理办法》,该办法为优化数字经济发展环境奠定了制度基础,对促进企业创新同样具有促进作用,上海市作为第一批试点城市效果尤为突出。为排除该政策干扰,本文剔除了来自上海市的样本,回归结果如表9第(3)列所示,仍旧与假设一致。2016年,贵州省获批建立我国首个国家级大数据综合试验区,对促进贵州省大数据基建整合和数据资源汇聚应用具有促进作用,鉴于此,本文剔除了贵州省的样本重新回归,结果如表9第(4)列所示,仍旧稳健。

(6)调整样本结构。鉴于本文的研究样本为上市公司,其子公司可能存在多地经营情况,进而受到多地政策冲击。为排除这一解释,本文还利用Python、Fassis 向量数据库等技术,从上市公司年报中获取到子公司经营情况,判断子公司与母公司的经营地一致性。稳健性检验中,进一步删除了上市公司本地子公司占比小于60%的样本。回归结果如表9第(5)列所示,由表9可知,*Reform* 在5%的水平上显著,结果稳健。

## 五、进一步分析

### 1. 细分效应检验:基于行业分类视角

数字经济包括数字产业化与产业数字化两条路径。按照《数字经济及其核心产业统计分类(2021)》标准,数字产业化对应数字经济核心产业<sup>①</sup>,主要包括计算机通信和其他电子设备制造业、电信广播电视和卫星传输服务、互联网和相关服务、软件和信息技术服务业等。而产业数字化指应用数字技术和数据资源为传统产业带来的产出增加和效率提升,是数字技术与实体经济的融合。依照上市公司行业分类指引(2012年版),本文样本共计包含18个一级行业代码和78个二级行业代码。以行业分类为视角,本文对不同行业性质样本分别进行数据回归。

首先,根据一级行业分类筛选样本<sup>②</sup>,回归结果如表10所示,其中仅“信息传输、软件和信息技术

<sup>①</sup> 按照《数字经济及其核心产业统计分类(2021)》标准,数字经济核心产业是指为产业数字化发展提供数字技术、产品、服务、基础设施和解决方案,以及完全依赖于数字技术、数据要素的各类经济活动。

<sup>②</sup> 为避免参与回归的样本总量过少造成结果偏差,保留样本总数大于200的九类一级行业。

术服务业”的政策效应观测项显著。制造业是国民经济的支柱性产业,依托技术创新驱动制造业提质增效,有利于中国经济实现高质量发展。鉴于此,本文还对制造业按照二级行业予以细分类,限于篇幅,仅对结果显著的二级制造业予以列示,如表 11 第(1)~(5)列所示。由表 11 可知,“农副食品加工业”“金属制品业”“计算机通信和其他电子设备制造业”的政策效应观测项显著为正,而“纺织业”和“通用设备制造业”出现负向显著结果。为细化研究结果,本文继续对“信息传输、软件和信息技术服务业”按照二级行业进行细分,由表 11 的第(6)~(8)列可知,“信息传输、软件和信息技术服务业”下属的“软件和信息技术服务业”呈现正向显著。

表 10 行业性质视角下的细分效应检验(一级行业分类)

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	文化、体育和娱乐业	科学研究和技术服务业	信息传输、软件和信息技术服务业	交通运输、仓储和邮政业	批发和零售业	建筑业	电力、热力及水生产和供应业	制造业	采矿业
<i>Reform</i>	-0.002 (-0.24)	-0.001 (-0.33)	0.010* (1.84)	0.001 (0.94)	-0.002 (-1.52)	-0.002 (-1.27)	0.001 (0.59)	0.000 (0.61)	0.004 (0.91)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制	控制
个体固定效应	是	是	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是	是	是
省份固定效应	是	是	是	是	是	是	是	是	是
常数项	0.160 (1.12)	0.064 (0.61)	0.169 (1.08)	0.237*** (3.53)	0.031 (1.04)	0.134** (2.07)	-0.036 (-0.70)	-0.005 (-0.22)	0.057 (0.71)
观测值	234	261	1605	340	421	424	462	15061	480
R <sup>2</sup>	0.324	0.275	0.244	0.351	0.346	0.513	0.149	0.231	0.317

表 11 行业性质视角下的细分效应检验(二级行业分类)

变量	制造业					信息传输、软件和信息技术服务业		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	农副食品加工业	纺织业	金属制品业	通用设备制造业	计算机、通信和其他电子设备制造业	电信、广播电视和卫星传输服务	互联网和相关服务	软件和信息技术服务业
<i>Reform</i>	0.003* (1.85)	-0.006** (-2.50)	0.003* (1.68)	-0.006** (-2.32)	0.006** (2.52)	0.003 (0.50)	0.007 (0.54)	0.013** (2.38)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制
个体固定效应	是	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是	是
省份固定效应	是	是	是	是	是	是	是	是
常数项	-0.055 (-0.83)	0.007 (0.09)	0.023 (0.13)	-0.026 (-0.33)	-0.096 (-1.27)	51.392 (1.03)	0.423* (1.78)	0.194 (1.04)
观测值	289	279	404	848	2197	93	347	1165
R <sup>2</sup>	0.278	0.424	0.338	0.339	0.369	0.721	0.297	0.309

综合上述回归结果,可以发现:第一,当前开放公共数据对企业创新的促进作用,主要集中在“信息传输、软件和信息技术服务业”这一以数据资源作为关键生产要素的数字经济核心产业,开

放公共数据对实现数字产业化已发挥了较为显著的促进作用,这与已有理论研究中提出的“数字产业的发展是数字经济发展的首要表现”相契合(史丹,2022)<sup>[16]</sup>。其原因可能在于,数字产业化作为数据生产要素支撑实体经济运行的基础层,在底层技术与生产过程中与数据生产要素的融合基础更好(倪克金和刘修岩,2021)<sup>[36]</sup>。第二,纺织业和通用设备制造业的回归结果并不理想,这可能与行业本身在要素密集度等方面呈现的特征有关。“通用设备制造业”等技术密集型行业企业对接数字技术的难度较大,且前期投入强度大、资产专用性强、转换成本高,数据生产要素对其产生的升级效应相对困难(朱秀梅和刘月,2021)<sup>[37]</sup>。“纺织业”生产模式相对传统,劳动密集程度高,依靠数据生产要素和数字技术进行产业升级的需求较低。王开科等(2020)<sup>[26]</sup>认为,数字技术在技术密集型和劳动密集型企业个体差异大,技术创新难以达到通用效果。第三,数据生产要素目前尚未完全显现出与实体经济结合的乘数效应,开放公共数据对产业数字化的促进作用仍需发力。从数据回归结果看,“农副食品加工业”和“金属制品业”正向显著,说明开放公共数据对智慧农业和智能制造等部分领域有积极影响。但总体来看,数据生产要素的内在价值未能在产业数字化领域全面得以释放。上述表现一方面与数据生产要素的开发基础有关,灾害、气象等基础环境数据与“农副食品加工业”进行季节性生产决策相关,应用场景关联度较高,涉及开发方案难度较低;另一方面,这与公共数据自身特征有关。公共数据与企业供给的商业性数据生产要素相比,颗粒度较大,主要聚焦在宏观层面。因此,数字经济时代,企业对数据生产要素的需求扩张,单纯依靠释放公共数据难以满足企业的创新需求。

## 2. 作用机制检验:基于创新资源和创新意愿视角

(1)资源补给机制:突破要素流通壁垒。理论部分提出,公共数据作为一类有价值的公共资源,被政府部门全权控制会产生寻租空间,诱发垄断性的资源攫取行为,从而导致资源配置的总体效率以及生产要素的利用效率下降。开放公共数据打通了公共数据的要素流通壁垒,有利于激发闲置要素内在价值,以禀赋升级为路径,推动企业实施技术创新。本文把这条影响机制总结为,开放公共数据对数据生产要素的资源补给机制。如果该机制存在,那么开放公共数据对企业创新的促进作用将在那些本身公共数据更匮乏的企业中更显著。政企纽带有助于企业获取资源和机会(Boubakri等,2008)<sup>[38]</sup>。公共数据掌握在公共部门手中,政企纽带为企业获取公共数据提供了优势。换言之,在公共数据实现开放共享前,政企纽带更紧密的企业更容易获取公共数据。

参考夏立军等(2011)<sup>[39]</sup>做法,本文首先设置政企纽带虚拟变量 $PC$ ,若企业董事长、总经理其中任意一人现任或曾任政府官员,则 $PC$ 取值为1,纳入有政企纽带组,否则为0,纳入无政企纽带组。同时,以 $PCLevel$ 表示政企纽带水平,按企业董事长、总经理任意一人曾任或现任政府机构职务最高级别分为四级,其中,无任职经历取0、任科级干部取1、任处级干部取2、任厅级干部取3、任部级干部取4,若取值为0和1,纳入政企纽带弱组,取值为2~4,纳入政企纽带强组。分组回归的结果如表12所示。由表12可知,政策效应观测项仅在无政企纽带和政企纽带弱的企业中显著,说明开放公共数据对数据生产要素的资源补给机制成立。

表 12 开放公共数据对数据生产要素的资源补给机制

变量	政企纽带虚拟变量		政企纽带强弱	
	有政企纽带	无政企纽带	政企纽带强	政企纽带弱
<i>Reform</i>	-0.000 (-0.35)	0.003*** (2.72)	-0.000 (-0.33)	0.002** (2.42)
控制变量	控制	控制	控制	控制
年份固定效应	是	是	是	是

续表 12

变量	政企纽带虚拟变量		政企纽带强弱	
	有政企纽带	无政企纽带	政企纽带强	政企纽带弱
公司固定效应	是	是	是	是
省份固定效应	是	是	是	是
常数项	0.040 (1.11)	-0.000 (-0.01)	0.041 (1.09)	-0.006 (-0.23)
观测值	6262	13956	5840	14379
R <sup>2</sup>	0.194	0.193	0.202	0.188
经验 P 值	b0-b1=0.03, p=0.000		b0-b1=0.02, p=0.000	

(2)动力激发机制:激励企业创新意愿。Zaltman 等(1973)<sup>[40]</sup>将创新意愿定义为企业进行创新的主观倾向。受收益滞后性强、失败率高等特点影响,企业创新延长了预期收益的实现时间,扩大了实施风险,容易导致企业创新意愿不足。理论分析部分中提出,开放公共数据强化了管理者竞争意识,降低了创新决策的不确定性,在一定程度上能够抵偿代理成本并激发企业创新意愿。本文把这条影响机制总结为,开放公共数据对企业创新意愿的动力激发机制。如果该机制存在,那么开放公共数据对企业创新的促进作用将在那些本身创新意愿较低的企业中更显著。公司治理理论认为,股权激励使管理者拥有更多经营者剩余索取权,从而提高管理者对股票价格波动的敏感程度,激发对未来绩效的关注与风险承担能力,转而做出长期性的战略发展规划(夏晗,2022)<sup>[5]</sup>。工资、奖金等薪酬激励具有较为明显的短期属性,对管理者采取薪酬激励能够在一定程度上抑制委托代理矛盾,并降低 CEO 的风险规避程度,鼓励管理者从事风险性投资。鉴于此,企业本身的股权激励水平越低或薪酬激励水平越低,企业创新意愿就越低,如果开放公共数据对创新意愿的动力激发机制存在,那么开放公共数据对企业创新的促进作用将在这类样本中更显著。

本文参考冯根福和温军(2008)<sup>[41]</sup>以及袁知柱等(2014)<sup>[42]</sup>做法,分别以高管持股比例即高管持股数占公司总股数的比例作为股权激励水平的代理变量,以及薪酬激励水平即前三位高管薪酬总额的自然对数作为薪酬激励水平的代理变量,并以上述变量的年度行业均值作为标准进行分组回归,结果如表 13 所示。由表 13 可知,政策效应观测项仅在股权激励水平低的企业中显著,数据生产要素对企业创新意愿不足的动力激发机制得证。但薪酬激励组却未呈现出差异,这说明,开放公共数据对企业创新意愿的激励作用,主要体现为针对长期激励不足的补偿。这是因为,公共数据作为一种新生产要素对市场及企业的改造具有系统性(肖土盛等,2022)<sup>[21]</sup>。刘淑春等(2021)<sup>[43]</sup>认为,在数据生产要素驱动企业实现创新转型时,往往需要经历一个漫长的过程,这种时滞性会形成一个“阵痛期”,在这个过程中,企业的用工模式、组织架构及管理流程都会随之发生改变(肖土盛等,2022)<sup>[21]</sup>。这使开放公共数据对创新意愿的激励作用与股权激励更相似,即针对战略发展规划实施的长期性激励。

表 13 数据生产要素对创新意愿的动力激发机制检验

变量	长期激励		短期激励	
	股权激励水平高	股权激励水平低	薪酬激励水平高	薪酬激励水平低
<i>Reform</i>	0.000 (0.15)	0.002** (2.43)	0.001 (1.10)	0.001 (1.44)
控制变量	控制	控制	控制	控制
年份固定效应	是	是	是	是
公司固定效应	是	是	是	是

续表 13

变量	长期激励		短期激励	
	股权激励水平高	股权激励水平低	薪酬激励水平高	薪酬激励水平低
省份固定效应	是	是	是	是
常数项	0.116*** (3.48)	0.030 (1.06)	0.058** (1.98)	-0.004 (-0.12)
观测值	9655	10564	9888	10331
R <sup>2</sup>	0.168	0.203	0.200	0.171
经验 P 值	b0-b1=0.02, p=0.000		b0-b1=0.00, p=0.474	

### 3. 异质性分析: 基于企业既有要素水平视角

大量观点认为,数据生产要素本身难以独立地创造价值。熊彼特创新理论提出创新的“深度模式”,认为创新主要来源于在位企业的技术累积性,因而先期进入的在位企业具有创新能力优势。李伟(2009)<sup>[44]</sup>认为,随着我国在技术转型过程中逐渐形成自主创新能力,此时本身拥有更多生产要素的大规模企业的创新能力更强。同时,从互补性资产的视角来看,企业在创新和市场化过程中,几乎都需要结合其他的能力和资产(Teece, 1986)<sup>[45]</sup>。尽管各省(自治区/直辖市)公共数据开放平台释放了大量可用于创新的数据生产要素,但部分企业由于资本体量、数据技术和市场融合经验等互补性资产储备不足,最终难以使数据价值与市场价值有效对接,从而抑制企业创新活力。因此,对那些在位累积性能力越强、互补性资产充足的企业而言,开放公共数据对企业创新的促进作用会更显著。

于立和王建林(2020)<sup>[1]</sup>提出,数字经济中的第一生产要素是数据,而资本、管理和技术是现代商业经济中其他重要生产要素。鉴于此,本文选取资本、管理和技术要素,针对开放公共数据对企业创新的关系进行异质性分析。其他生产要素的度量方式包括:第一,资本要素的度量。参考聂辉华等(2008)<sup>[46]</sup>,本文使用企业资产总额的自然对数度量企业拥有的资本要素情况,以年度行业平均数为标准进行分组,较高者纳入资本要素水平较高组,反之纳入较低组。第二,管理要素的度量。参考杨小娟等(2022)<sup>[47]</sup>做法,选用迪博(DIB)内部控制与风险管理数据库中“内部控制指数”的对数值进行衡量。本文以年度行业平均数为标准进行分组,较高者纳入管理要素水平较高组,反之纳入较低组。第三,技术要素的度量。根据企业是否为高科技行业度量企业拥有的技术要素情况,参考彭红星和毛新述(2017)<sup>[48]</sup>做法,制造业、信息传输、软件和信息技术服务业、科学研究和技术服务业涉及三个大类、19个小类的企业纳入高科技行业,其余纳入非高科技行业。回归结果如表 14 所示,发现政策效应观测项仅在企业规模较大组和高科技行业组,即资本及技术要素基础较好的企业中显著。但针对管理要素的考察中,内部控制较高组并未呈现显著,说明在企业利用数据生产要素进行创新的过程中,资本和技术要素相较管理要素更重要。

表 14 企业既有生产要素水平视角下的异质性分析

变量	资本要素水平		管理要素水平		技术要素水平	
	企业规模较大组	企业规模较小组	内部控制水平较高	内部控制水平较低	高科技行业	非高科技行业
<i>Reform</i>	0.003** (2.46)	0.001 (0.82)	0.001 (1.24)	0.000 (0.15)	0.002* (1.76)	0.000 (0.58)
控制变量	控制	控制	控制	控制	控制	控制
个体固定效应	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
省份固定效应	是	是	是	是	是	是

续表 14

变量	资本要素水平		管理要素水平		技术要素水平	
	企业规模较大组	企业规模较小组	内部控制水平较高	内部控制水平较低	高科技行业	非高科技行业
常数项	0.008 (0.21)	0.033 (1.18)	0.028 (1.00)	-0.002 (-0.07)	0.018 (0.89)	0.080*** (5.74)
观测值	9717	10502	10761	9458	13718	6501
R <sup>2</sup>	0.218	0.147	0.165	0.199	0.223	0.148
经验 P 值	b0-b1=-0.002, p= 0.012		b0-b1=-0.001, p=0.106		b0-b1=-0.001, p= 0.072	

## 六、研究结论与政策建议

有序释放数据生产要素是促进数据生产要素价值挖掘的前提。没有数据生产要素的投入,数字经济就失去了基本材料,难以推进后续加工并实现创新产品的转化。本文以各省(自治区/直辖市)陆续开通并运营公共数据开放平台为契机,利用多期双重差分模型,证实了开放公共数据与企业创新的关系。核心结论包括:第一,开放公共数据对企业创新总体上存在促进效应,但通过基于行业性质的细分检验发现,上述关系主要集中在“信息传输、软件和信息技术服务业”这一以数据资源作为关键生产要素的数字经济核心产业上,开放公共数据对实现数字产业化能够发挥显著促进作用;受行业特质影响,对“通用设备制造业”等技术密集型企业 and “纺织业”等劳动密集型企业而言,开放公共数据对企业创新的影响效果欠佳;尽管实证结果为开放公共数据对智慧农业和智能制造等少量领域的积极影响提供了佐证,但总的来说,目前开放公共数据仍未全面释放与实体经济相结合的乘数效应,对产业数字化的促进作用仍需发力。第二,数据质量和数据性质在现阶段是影响开放公共数据与企业创新关系的重要因素。从数据质量看,数据传输效率是影响开放公共数据与企业创新关系的关键因素,但数据总量和数据多样性在现阶段对上述关系影响效果不大,这说明单纯依靠量上的提升无法对创新产生实质性的影响,有价值的原材料才能真正符合生产要素本质;而当前企业对关联性应用场景的拓展不足,范围经济效应有待进一步深化。从数据性质看,当前释放文化教育和公共服务等性质的公共数据,对开放公共数据与企业创新的关系具有正向调节作用。第三,开放公共数据对企业创新的促进作用依托资源补给机制和动力激发机制为双向路径发挥作用。资源补给机制表现在,开放公共数据打通了公共数据的要素流通壁垒,引导公共数据作为重要生产要素流向企业;动力激发机制表现在,随着公共数据应用场景与方案的不断拓展,开放公共数据强化了管理者竞争意识,降低了创新决策中的不确定性,进而激发企业创新意愿。第四,企业的既有要素水平对开放公共数据与企业创新的关系具有异质性影响。对于企业规模较大和高科技行业的企业而言,开放公共数据对企业创新的促进作用更显著。由此提出如下政策启示:

第一,政府层面。从作用机制上看,一方面,要充分发挥开放公共数据对数据生产要素的资源补给机制,逐步扩大公共数据开放范围,建成并完善国家级公共数据开放平台。在顶层设计上,国家发展和改革委员会应以国家数据局为牵头单位,厘清各部门在数据管理及共享上的义务与权利,围绕数据价值评估、数据产权公证、数据隐私保护等方面统筹开展制度建设;在平台开发上,各级大数据工作主管部门应以资源整合为路径,率先打通政府机构和公共部门内部存在的数据生产要素流通壁垒,在建立跨部门、跨层级、跨省份的全国性公共数据内部共享平台基础上,进一步与面向社会公众的公共数据开放平台实现前端整合。另一方面,要充分发挥开放公共数据对企业创新意愿的动力补给机制,依托公共数据开放平台扩大成果辐射范围,营造开放竞争的数创环境。通过组织竞赛、购买服务、授权试点和数创沙龙等方式,引导企业、科研机构等多元主体利用公共

数据开展科学研究、咨询服务和产品开发等。围绕数据调用场景、服务对象和商业模式等内容,制定公共数据创新成果公开标准,定期组织项目申报与评审,对创新模式好、溢出效应显著、可复制性强的公共数据优秀创新成果进行遴选、发布与推广。与此同时,根据本文在数据质量和数据性质等维度的研究结论,各级大数据工作主管部门应遵循质量先行的原则,广泛征集公共数据开放需求,有侧重地优先开放社会需求迫切、商业潜力显著和民生诉求紧密相关的高价值公共数据,以应用效果和应用场景为牵引,突出公共数据的增值属性。同时,聚焦公共数据开放平台,通过完善分类检索导航、补充用户操作指南、增加 API 接口数量、提升数据可视化程度等方式,降低数据获取、分析与处理难度,提升数据资源处理能力,打造高质量的公共数据流通枢纽。

第二,企业层面。当前开放公共数据对促进数字产业化领域的创新作用明显,但依托产业数字化助力实体经济创新发展效果有限。鉴于此,首先,要保持企业在数字产业化领域对公共数据创新应用的先发优势,瞄准网络通信、集成电路、软件开发等数字经济核心产业,充分挖掘公共数据与战略性前瞻领域融合应用的细分市场,推动企业实现跨界创新。随着信用、教育、医疗、社会保障、文化、旅游等一系列既符合民生需求又契合市场需求的公共数据不断释放,企业应积极突出共享经济的开放性与生态性,发挥共赢与普惠作用,推动数字金融、互联网医疗、网络销售、自媒体娱乐等数字要素驱动业实现快速发展。其次,要激发企业在产业数字化领域对公共数据创新应用的巨大潜能,以实施数字化转型与智能化改造为路径,加快推动产业数字化进程。结合市场、顾客及产品的长尾需求,以新产业、新业态和新模式为突破口,推动传统产业实现智慧转型,挖掘企业数字化转型空间。强化数字化管理思维,提升管理层及普通员工的数字处理技能和数据管理能力,汇聚形成自有数据、公共数据和商业数据的多维网络,通过打造一体化数字平台,形成依托数据生产要素的决策驱动力,提升企业运行效率和产业链协同效率。最后,结合异质性分析结果,企业要积极研判自身的规模体量,制定科学的数字化转型路径与合理的投资计划。通过人才引进、专利引进、战略并购等方式,强化支撑数据创新的技术优势。积极抓住开放公共数据的改革契机,以原始业务为中心形成辐射效应,拓展一系列“公共数据+”应用场景,为企业长远发展寻找生产、经营和管理新方向。

## 参考文献

- [1]于立,王建林.生产要素理论新论——兼论数据要素的共性和特性[J].北京:经济与管理研究,2020,(4):62-73.
- [2]Veldkamp, L. Slow Boom Sudden Crash[J]. Journal of Economic Theory, 2005, 124, (2): 230-257.
- [3]Janssen, K. The Influence of the PSI Directive on Open Government Data: An Overview of Recent Developments[J]. Government Information Quarterly, 2011, 28, (4): 446-456.
- [4]Thorhildur, J., A. Michel, and B. A. Niels. Data-Driven Innovation through Open Government Data[J]. Journal of Theoretical & Applied Electronic Commerce Research, 2014, 9, (2): 100-120.
- [5]夏晗. 高管经历跨界、管理者激励与企业创新[J]. 北京: 科研管理, 2022, (2): 193-201.
- [6]林毅夫,张鹏飞. 适宜技术、技术选择和发展中国的经济增长[J]. 北京: 经济学(季刊), 2006, (3): 985-1006.
- [7]卡萝塔·佩蕾兹. 技术革命与金融资本[M]. 北京: 中国人民大学出版社, 2007.
- [8]胡业飞,孙华俊. 政府信息公开与数据开放的关联及治理逻辑辨析——基于“政府-市场-社会”关系变迁视角[J]. 北京: 中国行政管理, 2021, (2): 31-39.
- [9]黄曠琳,秦淑悦,张雨朦. 数字经济如何驱动制造业升级[J]. 北京: 经济管理, 2022, (4): 80-97.
- [10]Morton, F. S., P. Bouvier, A. Ezrachi, et al. Committee for the Study of Digital Platforms: Market Structure and Antitrust Subcommittee Report, Stigler Center for the Study of the Economy and the State[R]. University of Chicago Booth School of Business, 2019.
- [11]Bouncken, R. B., S. Kraus, and Roig-Tierno. Knowledge and Innovation-based Business Models for Future Growth: Digitalized Business Models and Portfolio Considerations[J]. Review of Managerial Science, 2019, 15, (1): 1-14.
- [12]Ciampi, F., S. Demi, A. Magrini, G. Marzi, and A. Papa. Exploring the Impact of Big Data Analytics Capabilities on Business Model Innovation: The Mediating Role of Entrepreneurial Orientation[J]. Journal of Business Research, 2021, 123: 1-13.

- [13] Basu, S., and D.N. Weil. Appropriate Technology and Growth[J]. Quarterly Journal of Economics, 1998, 113, (4): 1025-1054.
- [14] 约瑟夫·熊彼特. 资本主义, 社会主义与民主: 珍藏本[M]. 北京: 商务印书馆, 2009.
- [15] Einav, L., and J. Levin. Economics in the Age of Big Data[J]. Science, 2014, 346, (6210): 1-6.
- [16] 史丹. 数字经济条件下产业发展趋势的演变[J]. 北京: 中国工业经济, 2022, (11): 26-42.
- [17] Ciriello, R.F., A. Richter, and G. Schwabe. Digital Innovation[J]. Business and Information Systems Engineering, 2018, 60, (6): 563-569.
- [18] 唐要家, 王钰, 唐春晖. 数字经济、市场结构与创新绩效[J]. 北京: 中国工业经济, 2022, (10): 62-80.
- [19] Guelloc, D., and C. Paunov. Digital Innovation and the Distribution of Income[R]. NBER Working Papers, 2017, No. 23987.
- [20] Hagiu, A., and J. Wright. Multi-sided Platforms International[J]. Journal of Industrial Organization, 2015, (43): 162-174.
- [21] 肖土盛, 吴雨珊, 亓文韬. 数字化的翅膀能否助力企业高质量发展——来自企业创新的经验证据[J]. 北京: 经济管理, 2022, (5): 41-62.
- [22] 王超贤, 张伟东, 颜蒙. 数据越多越好吗——对数据要素报酬性质的跨学科分析[J]. 北京: 中国工业经济, 2022, (7): 44-64.
- [23] 林志杰, 孟政炫. 数据生产要素的结合机制——互补性资产视角[J]. 北京交通大学学报(社会科学版), 2021, (2): 28-38.
- [24] Nambisan, S., K. Lyytinen, A. Majchrzak, and M. Song. Digital Innovation Management: Reinventing Innovation Management Research in a Digital World[J]. MIS Quarterly, 2017, 41, (1): 223-238.
- [25] 刘玉奇, 王强. 数字化视角下的数据生产要素与资源配置重构研究——新零售与数字化转型[J]. 北京: 商业经济研究, 2019, (16): 5-7.
- [26] 王开科, 吴国兵, 章贵军. 数字经济发展改善了生产效率吗[J]. 成都: 经济学家, 2020, (10): 24-34.
- [27] 刘淑春. 中国数字经济高质量发展的靶向路径与政策供给[J]. 成都: 经济学家, 2019, (6): 52-61.
- [28] 姜松, 孙玉鑫. 数字经济对实体经济影响效应的实证研究[J]. 北京: 科研管理, 2020, (5): 32-39.
- [29] 徐悦, 刘运国, 蔡贵龙. 高管薪酬粘性与企业创新[J]. 北京: 会计研究, 2018, (7): 43-49.
- [30] 郭玥. 政府创新补助的信号传递机制与企业创新[J]. 北京: 中国工业经济, 2018, (9): 98-116.
- [31] 解维敏, 唐清泉, 陆姗姗. 政府R&D资助, 企业R&D支出与自主创新——来自中国上市公司的经验证据[J]. 北京: 金融研究, 2009, (6): 86-99.
- [32] 李春涛, 同续文, 宋敏, 杨威. 金融科技与企业创新——新三板上市公司的证据[J]. 北京: 中国工业经济, 2020, (1): 81-98.
- [33] 鞠晓生, 卢荻, 虞义华. 融资约束、营运资本管理与企业创新可持续性[J]. 北京: 经济研究, 2013, (1): 4-16.
- [34] 陶锋, 朱盼, 邱楚芝, 王欣然. 数字技术创新对企业市场价值的影响研究[J]. 北京: 数量经济技术经济研究, 2023, (5): 68-91.
- [35] Beck T., and R.L. Levkov. Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States[J]. Journal of Finance, 2010, (5): 1637-1667.
- [36] 倪克金, 刘修岩. 数字化转型与企业成长: 理论逻辑与中国实践[J]. 北京: 经济管理, 2021, (12): 79-97.
- [37] 朱秀梅, 刘月. 企业数智转型能力形成机理——基于海尔集团“知行合一”的单案例研究[J]. 北京: 经济管理, 2021, (12): 98-114.
- [38] Boubakri, N., J. Cosset, and W. Saffar. Political Connections of Newly Privatized Firms[J]. Journal of Corporate Finance, 2008, 14, (5): 654-673.
- [39] 夏立军, 陆铭, 余为政. 政企纽带与跨省投资——来自中国上市公司的经验证据[J]. 北京: 管理世界, 2011, (7): 128-140.
- [40] Zaltman, G., R. Dunaev, and J. Holbek. Innovations and Organizations[M]. New York: John Wiley, 1973.
- [41] 冯根福, 温军. 中国上市公司治理与企业技术创新关系的实证分析[J]. 北京: 中国工业经济, 2008, (7): 91-101.
- [42] 袁知柱, 郝文瀚, 王泽桑. 管理层激励对企业应计与真实盈余管理行为影响的实证研究[J]. 北京: 管理评论, 2014, (10): 181-196.
- [43] 刘淑春, 同津臣, 张思雪, 林汉川. 企业管理数字化变革能提升投入产出效率吗[J]. 北京: 管理世界, 2021, (5): 170-190, 13.
- [44] 李伟. 产业演进中的技术创新与市场结构关系——兼论熊彼特假说的中国解释[J]. 北京: 科研管理, 2009, (6): 39-47.
- [45] Teece, D.J. Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy[J]. Research Policy, 1986, 15, (6): 285-305.
- [46] 聂辉华, 谭松涛, 王宇锋. 创新、企业规模和市场竞争: 基于中国企业层面的面板数据分析[J]. 北京: 世界经济, 2008, (7): 57-66.
- [47] 杨小娟, 李方晗, 赵艺. 连锁董事的内部控制溢出效应研究——基于内部控制缺陷的视角[J]. 北京: 审计研究, 2022, (3): 117-128.
- [48] 彭红星, 毛新述. 政府创新补贴、公司高管背景与研发投入——来自我国高科技行业的经验证据[J]. 北京: 财贸经济, 2017, (3): 147-161.

# Open Public Data and Enterprise Innovation from the Perspective of Data Production Factors: A Quasi-Natural Experiment of Establishing an Open Platform for Public Data

CHEN Yan-li<sup>1</sup>, JIANG Qi<sup>2</sup>

(1.School of Accounting/China Internal Control Research Center,Dongbei University of Finance and Economics, Dalian,Liaoning,110625,China;

2.School of Political Science and Public Administration,Southwest University of Political Science and Law, Chongqing,401120,China)

**Abstract:** As a strategic factor of production in the era of digital economy, data is leading the innovation revolution and technological change. Public data refers to the data resources generated, collected, recorded and saved by public departments such as state organs, institutions and scientific research institutions in the process of performing public functions. In 2016, China's National Reform teleconference proposed that "In the era of big data, essential information for people's daily needs such as navigation, weather, housing, medical care, employment frequently originates from the open data made available by government sources; The strategic thinking, layout planning, and implementation plans required for industrial development often rely on the mining, restructuring, and mixing of government information and data..." Although the innovative value of public data as an emerging factor of production has been widely recognized, its monopolistic nature in terms of sources and the obstacle of factor circulation influenced by the characteristics of high collection cost, centralized subject and strong administrative attribute, may lead to an overall decline in the efficiency of resource allocation. It is crucial to break through the barriers to the circulation of public data, promote the flow of public data as an innovation resource to firms, and encourage and guide firms to utilize public data for innovation decisions and outcome transformation.

This paper employs a staggered difference-in-difference model based on the quasi-natural experiment of the public data open platform successively opened and operated by various provinces (autonomous regions/municipalities directly under the Central Government) in China, which confirms that the open access to public data can promote firm innovation. The above relationship is based on the resource supply mechanism of data production factors and the dynamic stimulation mechanism of innovation willingness as a two-way path. Based on the segmentation examination of the industry characteristics, it is found that the current release of data production factors has an obvious effect on the realization of digital industrialization, but the effect of relying on industrial digitalization to help the real economy is limited. Further, this paper finds that data transmission efficiency is the key factor affecting the above relationship, and the open cultural education and public service data production factors are more effective at this stage. At the same time, based on the heterogeneity analysis of existing factor levels, it is found that the relationship is more significant for firms with stronger presence accumulation of production factors such as technology and capital.

The innovation of this paper includes the following aspects: First, it adds the theoretical analysis and empirical evidence about the economic value of data production factors. This paper constructs the theoretical analysis framework of the relationship between open public data and firm innovation for the first time, and analyzes the influencing factors, mechanism and heterogeneity of the relationship based on the industry classification, data quality, data nature and other dimensions of the digital economy, providing a theoretical reference for confirming the actual economic value of data production factors. Second, it enriches the research perspective and conclusions of the factors affecting firm innovation. As for the research in the same field, there is still a lack of literature that systematically discusses the relationship between open public data and firm innovation by including open public data into the category of influencing factors of firm innovation. Data production factors and data innovation are significantly different from other production factors and traditional innovation forms. The analytical framework of traditional firm innovation needs to be improved in combination with the background of the digital economy era to meet the changing requirements of economic growth characteristics in the digital era. To some extent, this paper has made incremental contributions in this regard. The research conclusions are not only conducive to boosting the government's confidence in the implementation of public data opening, providing policy references for promoting the establishment of data production factor market trading system and continuously improving the construction of public data open platform, but also provide a feasible path for implementing innovation-driven strategy and stimulating the innovation vitality of firms in the era of digital economy.

**Key Words:** data production factor; public data; enterprise innovation; open public data

**JEL Classification:** D22, G39

**DOI:** 10.19616/j.cnki.bmj.2024.01.002

(责任编辑:张任之)